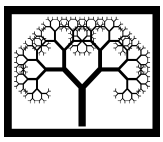
Paper 66



Proceedings of the Tenth International Conference
on Computational Structures Technology,
B.H.V. Topping, J.M. Adam, F.J. Pallarés,
R. Bru and M.L. Romero, (Editors),
Civil-Comp Press, Stirlingshire, Scotland

Maximum Likelihood Estimation of Modal Parameters in Structures Using the Expectation Maximization Algorithm

F.J. Cara[1], J. Carpio[1], J. Juan[1] and E. Alarcon[2]
[1] Department of Organization Engineering,
Business Administration and Statistics
[2] Department of Structural Mechanics and Industrial Constructions
Polytechnical University of Madrid, Spain

Abstract

This paper presents a time-domain stochastic system identification method based on Maximum Likelihood Estimation (MLE) with the Expectation Maximization (EM) algorithm. The effectiveness of this structural identification method is evaluated through numerical simulation in the context of the ASCE benchmark problem on structural health monitoring. Modal parameters (eigenfrequencies, damping ratios and mode shapes) of the benchmark structure have been estimated using both Stochastic Subspace Identification (SSI) method and the proposed MLE+EM method. The numerical results show that the proposed method estimates more accurate modal parameters than SSI in the presence of 10% measurement noise. Finally, adventages and disadventages of the method have been discussed.

Keywords: system identification in structures, state space models, Kalman filter, stochastic subspace methods, modal analysis, benchmark problems.

1 Introduction

The application of system identification to vibrating structures consist in identifying a modal model (eigenfrequencies, damping ratios and mode shapes) from vibration data. Classically, a measurable input is applied to the system and the output is measured. From these experimental data, a system model can be obtained by a variety of parameter estimation methods, and it is known as experimental modal analysis. However, cases exist where it is practically impossible to measure the excitation and the outputs are the only information that is passed to the system identification algorithms. In these cases the deterministic knowledge of the input is replaced by the assumption that the input is a realization of a stochastic process (white noise), and it is known as stochastic system identification (the terms output-only modal analysis and operational

modal analysis are used as well).

Parametric structural identification methods involve the use of mathematical models to represent structural system behavior in either time or frequency domain. The benefits of using parametric models for structural identification include their direct relationship with physically meaningful quantities such as stiffness and mass, improved accuracy and resolution, and their suitability for analysis, prediction, fault diagnosis and control.

Popular time domain parametric models used for structural identification purposes include: ARX models, ARMAX models, state space models, etc. Many identification algorithms are available to estimate the parameters of such parametric models, e.g. prediction error method (PEM), least squares estimation (LSE), maximum likelihood algorithm (MLA), eigensystem realization algorithm (ERA) and stochastic subspace identification method (SSI).

This paper presents a time-domain stochastic system identification method based on Maximum Likelihood Estimation (MLE) with the Expectation Maximization (EM) algorithm. The effectiveness of this structural identification method is evaluated through numerical simulation in the context of the ASCE benchmark problem on structural health monitoring [7]. Modal parameters (eigenfrequencies, damping ratios and mode shapes) of the benchmark structure (see Figure 2) have been estimated using both Stochastic Subspace Identification (SSI) method and the proposed MLE+EM method.

SSI identification method is a well known method and computes accurate estimates of the modal parameters ([9], [10]), and for this reason it has been used for comparison. The principles of the SSI method have been introduced in the paper and next the proposed Maximum Likelihood Estimation with Expectation Maximization algorithm has been explained in detail. Finally, the results obtained with both methods are compared.

2 State space model

A vibrating structure can be represented by a discrete-time stochastic state-space model given as:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \\ y_k &= Cx_k + Du_k + v_k \end{aligned} \tag{1}$$

where

k denotes the sampling instant ($t = k\Delta t$, with constant sampling time Δt);

$y_k \in \mathbb{R}^l$ is the measured output vector;

$u_k \in \mathbb{R}^m$ is the measured input vector;

$x_k \in \mathbb{R}^n$ is the discrete state vector;

$w_k \in \mathbb{R}^n$ is the noise due to disturbances and modeling inaccuracies;

$v_k \in \mathbb{R}^l$ is the measurement noise due to sensor inaccuracy

$A \in \mathbb{R}^{n \times n}$ is the transition state matrix describing the dynamics of the system (as characterised by its eigenvalues);

$B \in \mathbb{R}^{n \times m}$ is the input matrix;

$C \in \mathbb{R}^{l \times n}$ is the output matrix, which is describing how the internal state is transferred to the the output measurements y_k;

$D \in \mathbb{R}^{l \times m}$ is the direct transmission matrix;

The noise vectors comprise unmeasurable vector signals assumed to be zero-mean with covariance matrices

$$E\left[\begin{pmatrix} w_p \\ v_p \end{pmatrix} \begin{pmatrix} w_p^T & v_p^T \end{pmatrix}\right] = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{pq} \tag{2}$$

where E is the expected value operator and δ_{pq} is the Kronecker delta.

In the case of ambient vibration testing, only the responses of the structure y_k are measured, while the input sequence u_k remains unmeasured. Equation (1) results now in a purely stochastic system:

$$\begin{aligned} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k \end{aligned} \tag{3}$$

The input is now implicitly modeled by the noise terms w_k, v_k. However the white noise assumptions of these noise terms cannot be omitted and (2) remain still applicable in equation (3).

3 Parameter estimation methods

The system identification problem in the state space model defined in Equation (3) can be formulated as the determination of the order n and the corresponding system matrices A and C (up to within a similarity transformation) using the output measurements $\{y_1, y_2, \ldots, y_N\}$ available for N time steps. In the case of parametric system identification methods, the dynamic behavior of a system is described using mathematical models and mathematical relationships between the modal parameters and the estimated model parameters (A,C).

There are several techniques to realize system identification. In this work, the data-driven stochastic subspace identification method (SSI) has been used for comparison. In the following section the principles of the SSI identification method will be introduced and next the proposed Maximum Likelihood Estimation with Expectation Maximization algorithm will be explained in detail.

3.1 Stochastic subspace identification method for state space models

Subspace methods identify state-space models from (input and) output data by applying robust numerical techniques such as QR factorization, SVD and least squares. The first SSI algorithms can be found in [6], and a general overview of data-driven subspace identification (both deterministic and stochastic) is provided in [2]. A brief description of the method is included following.

Let us for a moment assume that not only is y_k measured, but also the sequence of state vectors x_k. Thus, with known y_k and x_k, the model (3) becomes a linear regression. To see this clearly, let:

$$Z_k = \begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} \in \mathbb{R}^{(n+l)\times 1} \qquad \Phi_k = x_k \in \mathbb{R}^{n\times 1}$$

$$\theta = \begin{bmatrix} A \\ C \end{bmatrix} \in \mathbb{R}^{(n+l)\times n} \qquad E_k = \begin{bmatrix} w_k \\ v_k \end{bmatrix} \in \mathbb{R}^{(n+l)\times 1}$$

Then, (3) can be rewritten as:

$$Z_k = \theta\Phi_k + E_k \tag{4}$$

From this, all the matrix elements in θ can be estimated by the simplest least squares method as follows. The criterion function is defined as:

$$V_N(\theta) = \frac{1}{N}\sum_{k=1}^{N}\left(Z_k - \theta\Phi_k\right)^2 \tag{5}$$

The least square estimate $\hat{\theta}$ is defined by minimization of $V_N(\theta)$. Analytically, setting the gradient of $V_N(\theta)$ with respect to θ to zero, yields:

$$\hat{\theta} = \left(\sum_{k=1}^{N} Z_k\Phi_k^T\right)\left(\sum_{k=1}^{N}\Phi_k\Phi_k^T\right)^{-1} \tag{6}$$

Moreover, the the residuals and its covariance matrices are given by:

$$\hat{E}_k = Z_k - \hat{\theta}\Phi_k \tag{7}$$

$$\begin{bmatrix} \hat{Q} & \hat{S} \\ \hat{S}^T & \hat{R} \end{bmatrix} = \frac{1}{N}\sum_{k=1}^{N}\hat{E}_k\hat{E}_k^T \tag{8}$$

Thus, knowing a sequence of state vectors x_k, the problem given by (4) is solved and the unknown matrices A and C are computed. Subspace methods are, in essence, numerical methods to construct a good estimate of a sequence of state vectors of the state space model from the measured output data.

In the following it is briefly explained how subspace methods work. First, the stochastic system (3) can be converted into a so-called forward innovation model by applying the Kalman filter:

$$\begin{aligned} x_{k+1} &= Ax_k + Ke_k \\ y_k &= Cx_k + e_k \end{aligned} \tag{9}$$

Then, a non-steady state Kalman filter state estimate $\hat{x}_k$ is defined by the following recursive formulae:

$$\hat{x}_k = A\hat{x}_{k-1} + K\left(y_{k-1} - C\hat{x}_{k-1}\right) \tag{10}$$

the Kalman filter state estimate can be written as [2]:

$$\hat{x}_k = L_k \begin{bmatrix} y_0 \\ y_1 \\ \cdots \\ y_{k-1} \end{bmatrix} \tag{11}$$

A linear combination of the past output measurements $y_0, \ldots, y_{k-1}$ ($L_k \in \mathbb{R}^{n\times(kl)}$), which allows for the definition of the Kalman filter state sequence of j states as:

$$\hat{X}_i = [\underbrace{\hat{x}_i \quad \hat{x}_{i+1} \quad \ldots \quad \hat{x}_{i+j-1}}_{j\ states}] = L_i \begin{bmatrix} y_0 & y_1 & \cdots & y_{j-1} \\ y_1 & y_2 & \cdots & y_{j-2} \\ \cdots & \cdots & \cdots & \cdots \\ y_{i-1} & y_i & \cdots & y_{i+j-2} \end{bmatrix} = L_i Y_p \tag{12}$$

where Y_p is the block Hankel matrix of past outputs.

$$Y_p = \begin{bmatrix} y_0 & y_1 & \cdots & y_{j-1} \\ y_1 & y_2 & \cdots & y_{j-2} \\ \cdots & \cdots & \cdots & \cdots \\ y_{i-1} & y_i & \cdots & y_{i+j-2} \end{bmatrix} \tag{13}$$

In other words, Y_p forms a row basis for the computation of the state sequence needed in (4). Nevertheless, subspace methods don't compute $\hat{X}_i$ directly from Y_p, but from a projection onto Y_p.

$$\hat{X}_i = \Gamma_i^{-1}\left[Y_f / Y_p\right] \tag{14}$$

where:

- Γ_i is the extended observability matrix.
- Y_f is the block Hankel matrix of future outputs, defined in a similar way like Y_p, but starting from i.

$$Y_f = \begin{bmatrix} y_i & y_{i+1} & \cdots & y_{i+j-1} \\ y_{i+1} & y_{i+2} & \cdots & y_{i+j} \\ \cdots & \cdots & \cdots & \cdots \\ y_{2i-1} & y_{2i} & \cdots & y_{2i+j-2} \end{bmatrix} \tag{15}$$

- $[Y_f/Y_p]$ is the orthogonal projection of Y_f onto Y_p. This projection is computed using LQ decomposition.

Suppose that the singular value decomposition of $[Y_f/Y_p]$ is given by $[Y_f/Y_p] = USV^T$ with $rank(S) = n$. Thus, the extended observability matrix can be taken as $\Gamma_i = US^{1/2}$. Hence, it follows that the state sequence estimate is given by

$$\hat{X}_i = S^{1/2}V^T \tag{16}$$

This is the subspace projection approach, that applied robust numerical techniques like LQ decomposition and singular value decomposition to Hankel matrices formed with outputs measurements only to estimate the matrices of the state space model (a detailed description of different algorithms which implement subspace identification can be found in [2] and [4]).

3.2 Proposed maximum likelihood method with EM algorithm for state space models

In this section is presented the proposed identification algorithm for estimating the parameters of the stochastic state space model given by (3), which is based on the maximum likelihood method. This method try to maximize the likelihood applying the iterative expectation maximization algorithm (EM). The proposed identification method starts computing the likelihood in the state space model:

Given N measurements of the outputs $Y_N = \{y_1, y_2, \ldots, y_N\}$, a vector θ is defined to represent the unknown parameters of the model (3):

$$\theta \stackrel{def}{=} (A, C, Q, R, \mu_0, \Sigma_0)$$

under the assumption that the initial state is normal, $x_0 \rightsquigarrow N(\mu_0, \Sigma_0)$. The likelihood is computed using the innovations $\epsilon_1, \epsilon_1, \ldots, \epsilon_1$, defined by [1]:

$$\epsilon_k = y_k - CX_k^{k-1} \tag{17}$$

The innovations form of the likelihood is obtained by noting the innovations are independent Gaussian random vectors with zero means and covariance matriz

$$\Sigma_k = CP_k^{k-1}C' + R \tag{18}$$

Hence, ignoring a constant, we may write the logarithm of the likelihood, $L_{Y_N}(\theta)$, as:

$$l_{Y_N}(\theta) = logL_{Y_N}(\theta) = -\frac{1}{2}\sum_{k=1}^{N} log\,|\Sigma_k(\theta)| - \frac{1}{2}\sum_{k=1}^{N} \epsilon_k(\theta)'\Sigma_k(\theta)^{-1}\epsilon_k(\theta) \tag{19}$$

[1] Given the output data for s time steps $Y_s = \{y_1, y_2, \ldots, y_s\}$ is defined:

$$x_k^s = \mathbf{E}\left[x_k|Y_s\right]$$

$$P_{k_1,k_2}^s = \mathbf{E}\left[(x_{k_1} - x_{k_1}^s)(x_{k_2} - x_{k_2}^s)^T|Y_s\right]$$

where $\mathbf{E}\left[*|*\right]$ is the conditional expected operator. When $k_1 = k_2 = k$ it will be written P_t^s.

where it has been emphasized the dependence of the innovations on the parameters θ. Of course, (19) is a highly nonlinear and complicated function of the unknown parameters. The usual procedure is to fix x_0 and then develop a set or recursions for the log likelihood function and its first two derivatives. Then, a Newton-Raphson algorithm can be used successively to update the parameter values until likelihood is maximized.

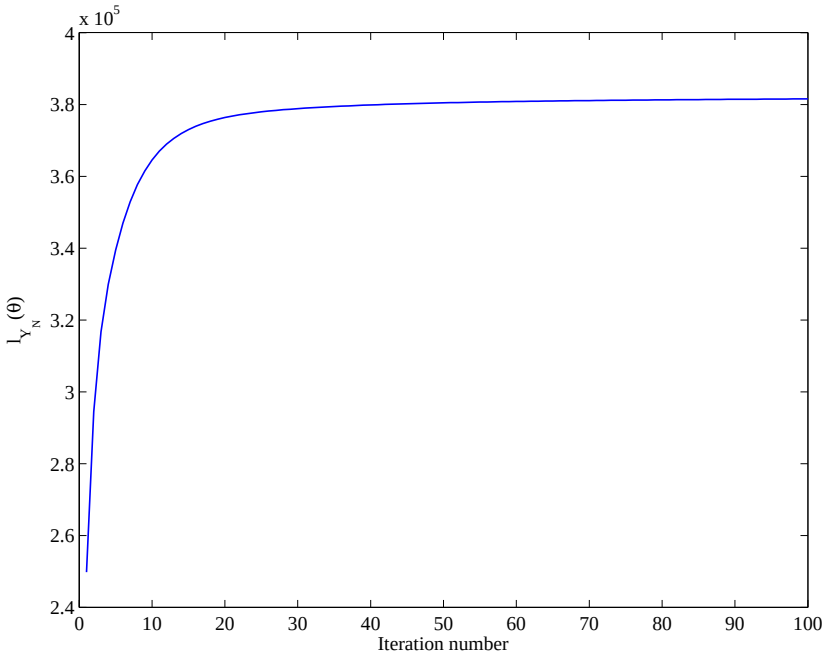


Figure 1: Likelihood $L_{Y_N}(\theta)$ of a simulated case (it corresponds to case 1, numerical example section)

In addition to Newton-Raphson, Shumway and Stoffer [3] presented a conceptually simpler estimation procedure based on the Expectation-Maximitation algorithm (EM). The EM algorithm is simple to apply since at each iteration the optimal solution for the unknown parameters can be obtained from explicit regression formulas.

The basic idea is that if we could observe the states $X_N = \{x_0, x_1, x_2, ..., x_N\}$, in addition to the observations, $Y_N = \{y_1, y_2, \ldots, y_N\}$, then we could consider the complete data $Z_N = \{X_N, Y_N\}$, with the joint density

$$f(Z_N|\theta) = f_{\mu_0,\Sigma_0}(x_0) \prod_{k=1}^{N} f_{A,Q}(x_k|x_{k-1}) \prod_{k=1}^{N} f_{C,R}(y_k|x_k) \tag{20}$$

where under the Gaussian assumption

$$f_{\mu_0,\Sigma_0}(x_0) = \frac{1}{(2\pi)^{n/2}\,|\Sigma_0|^{1/2}} \exp\{-\frac{1}{2}(x_0-\mu_0)^T\Sigma_0^{-1}(x_0-\mu_0)\} \tag{21}$$

$$f_{A,Q}(x_k|x_{k-1}) = \frac{1}{(2\pi)^{n/2}\,|Q|^{1/2}} \exp\{-\frac{1}{2}(x_k - Ax_{k-1})^T Q^{-1}(x_k - Ax_{k-1})\} \quad (22)$$

$$f_{C,R}(y_k|x_k) = \frac{1}{(2\pi)^{l/2}\,|R|^{1/2}} \exp\{-\frac{1}{2}(y_k - Cx_k)^T R^{-1}(y_k - Cx_k)\} \quad (23)$$

and the "complete" data likelihood is defined by $L_{X_N,Y_N}(\theta) = f(Z_N\,|\theta)$. If we had the complete data Z_N, the maximum likelihood estimators (MLEs) of θ would be easily obtained from $L_{X_N,Y_N}(\theta)$. The problem is more difficult than this, because we do not know X_N, and we have to estimate the parameters θ from just the observed information Y_N (the likelihood of θ given Y_N is $L_{Y_N}(\theta)$ and it is related with $L_{X_N,Y_N}(\theta)$).

The EM provides an iterative method for finding the MLEs of θ by succesively maximizing the conditional expectation of the complete likelihood $L_{X_N,Y_N}(\theta)$. The log-likelihood $l_{X_N,Y_N}(\theta) = \log L_{X_N,Y_N}(\theta)$ is preferred because information can be written as a sum of three uncoupled functions

$$l_{X_N,Y_N}(\theta) = -\frac{1}{2}[l_1(\mu_0, \Sigma_0) + l_2(A, Q) + l_3(C, R))]$$

where, ignoring constants

$$l_1(\mu_0, \Sigma_0) = \log|\mathbf{\Sigma}_0| + (x_0 - \mu_0)^T \mathbf{\Sigma}_0^{-1}(x_0 - \mu_0) \quad (24)$$

$$l_2(A, Q) = N \log|Q| + \sum_{k=1}^{N}(x_k - Ax_{k-1})^T Q^{-1}(x_k - Ax_{k-1}) \quad (25)$$

$$l_3(C, R) = N \log|R| + \sum_{k=1}^{N}(y_k - Cx_k)^T R^{-1}(y_k - Cx_k) \quad (26)$$

Each iteration of the EM algorithm consists of two steps. If θ_j denotes the estimated values of the parameter θ after j iterations, the first step (E step) of the next iteration $j+1$ is to compute

$$S(\theta|Y_N, \theta_j) = E[l_{X_N,Y_N}(\theta)|Y_N, \theta_j]. \quad (27)$$

$S(\theta|Y_N, \theta_j)$ is the key function of this method. The second step (M step) consists on maximizing $S(\theta|Y_N, \theta_j)$, what is equivalent to maximize the likelihood $L_{Y_N}(\theta)$ (see fig. 1).

3.2.1 E-step: computation of $S(\theta|Y_N, \theta_j)$

Note that given Y_N and θ_j the only terms that remain random (unknown) in (25) are the states, x_k. In this step, the algorithm compute the expected values of (25) respect to x_k. Given the value of the parameters θ for iteration j, the Kalman filter and smoother provides the following values for $k = 0, 1, \ldots, N$ (see Appendix A):

$$x_k^N = E[x_k|Y_N, \theta_j] \quad (28)$$

$$P_k^N = E[(x_k - x_k^N)(x_k - x_k^N)^T|Y_N, \theta_j] \quad (29)$$

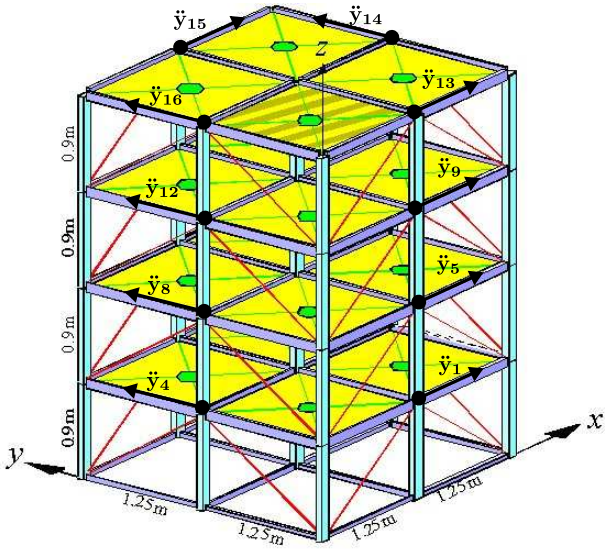


Figure 2: Diagram of the analytical model of the benchmark structure

$$P^N_{k,k-1} = E[(x_k - x^N_k)(x_{k-1} - x^N_{k-1})^T | Y_N, \theta_j] \tag{30}$$

and from them it is possible to compute

$$E[l_1(\mu_0, \Sigma_0)|Y_N, \theta_j] = \log|\Sigma_0| + \text{tr}\left\{\Sigma_0^{-1}\left[P^N_0 + (x^N_0 - \mu_0)(x^N_0 - \mu_0)\right]\right\} \tag{31}$$

$$E[l_2(A, Q)|Y_N, \theta_j] = N\log|Q| + \text{tr}\left\{Q^{-1}\left[S_{xx} - S_{xb}A^T - AS_{bx} + AS_{bb}A^T\right]\right\} \tag{32}$$

$$E[l_3(C, R)|Y_N, , \theta_j] = N\log|R| + \text{tr}\left\{R^{-1}\left[S_{yy} - S_{yx}C^T - CS_{xy} + CS_{xx}C^T\right]\right\} \tag{33}$$

where

$$S_{xx} = \sum_{k=1}^{N}\left(P^N_k + x^N_k(x^N_k)^T\right) \tag{34}$$

$$S_{xb} = \sum_{k=1}^{N}\left(P^N_{k,k-1} + x^N_k(x^N_{k-1})^T\right)) \tag{35}$$

$$S_{bb} = \sum_{k=1}^{N}\left(P^N_{k-1} + x^N_{k-1}(x^N_{k-1})^T\right)) \tag{36}$$

$$S_{yy} = \sum_{k=1}^{N}\left(y_k y_k^T\right)) \tag{37}$$

$$S_{yx} = \sum_{k=1}^{N}\left(y_k(x^N_{k-1})^T\right)) \tag{38}$$

The function $S(\theta|Y_N, \theta_j)$ is the sum of the three terms of Equations (31)-(33), and it depends on the parameters $\theta = (A, C, Q, R, \mu_0, \Sigma_0)$. In the next section, the values $\theta = \theta_{j+1}$ that maximize S are obtained.

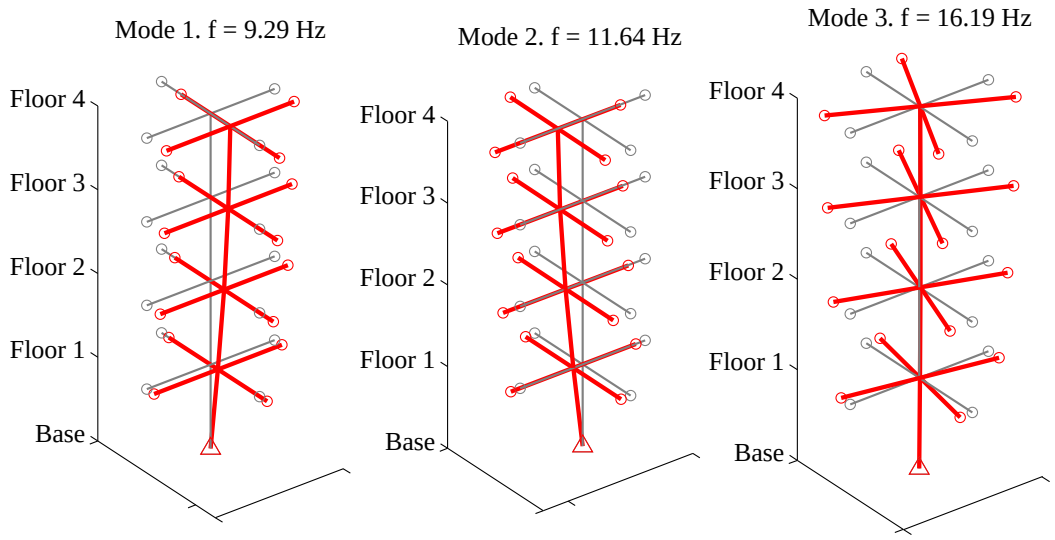


Figure 3: First 3 eigenfrequencies and mode shapes computed from 12 DOF matrices (exact values)

3.2.2 M-Step: maximization of $S(\theta|Y_N, \theta_j)$

Maximizing $S(\theta|Y_N, \theta_j)$ with respect of the parameters θ, at iteration j, constitutes the M-step and is analogous to the multivariate regression approach. This is the strong point of th EM algorithm, the maximum values are obtained from explicit formula.

The maximum of $E[l_1(\mu_0, \Sigma_0)|Y_n, \theta_j]$ is attained at

$$\hat{\mu}_0 = x_0^N \tag{39}$$

$$\hat{\Sigma}_0 = P_0^N \tag{40}$$

The estimation of A can be found equating to zero the derivative of $E[l_2(A, Q)|Y_N, \theta_j]$:

$$\hat{A} = S_{xb} S_{bb}^{-1} \tag{41}$$

and

$$\hat{Q} = \frac{1}{N}\left(S_{xx} - S_{xb}\hat{A}^T - \hat{A}S_{bx} + \hat{A}S_{bb}\hat{A}^T\right) \tag{42}$$

In a similar way from $E[l_3(C, R)|Y_n, \theta_j]$ the estimation of C and R are:

$$\hat{C} = S_{yx} S_{xx}^{-1} \tag{43}$$

$$\hat{R} = \frac{1}{N}\left(S_{yy} - S_{yx}\hat{C}^T - \hat{C}S_{xy} + \hat{C}S_{xx}\hat{C}^T\right) \tag{44}$$

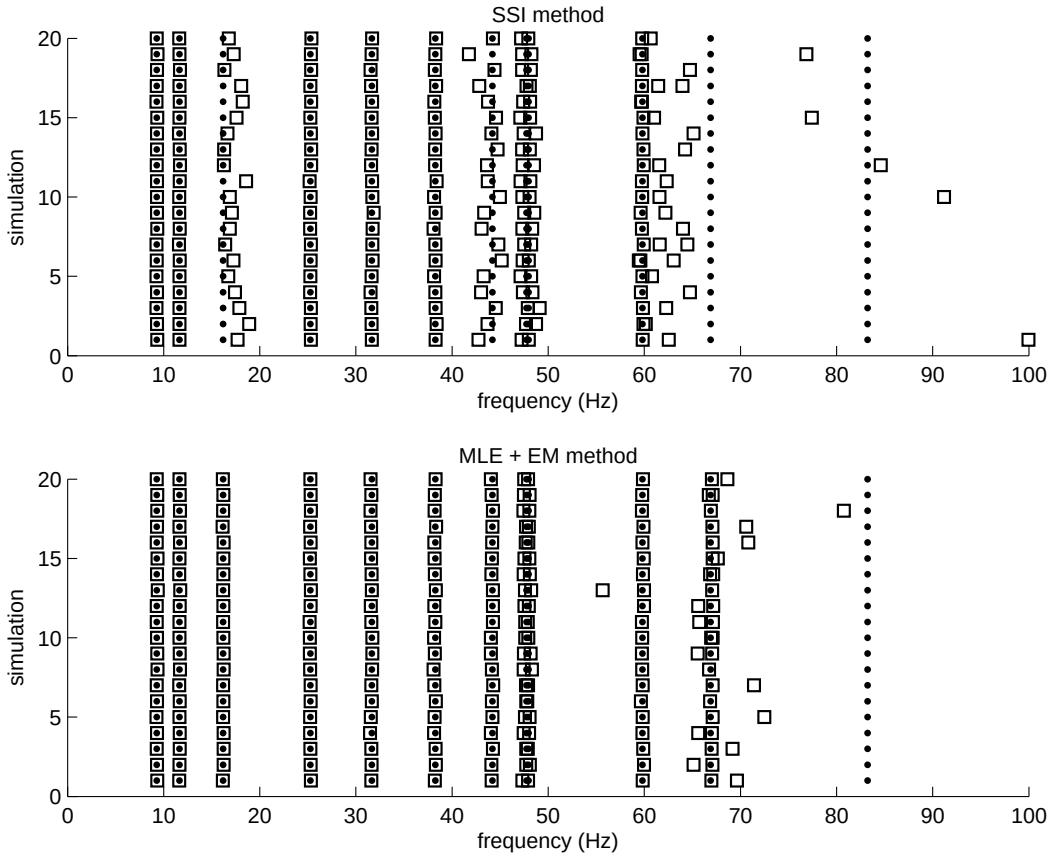


Figure 4: Eigenfrequency results corresponding to the first 20 simulated cases for the SSI and MLE+EM methods. The symbols "▪" and "□" denote exact and estimated eigenfrequency, respectively.

3.2.3 Overall procedure

The overall method can be summarized as an iterative procedure as follows:

1. Initialize the procedure by selecting starting values for the parameters $\theta^0 = (A, C, Q, R, \mu_0, \Sigma_0)$.

   On iteration j ($j = 1, 2, \ldots$)

2. Compute the incomplete-data likelihood, $L_{Y_N}(\theta^{(j-1)})$.

3. Perform the E-Step. Use Properties A.1, A.2 y A.3 to obtain the smoothed values x_k^N, P_k^N, and $P_{k,k-1}^N$, for $k = 1, 2, \ldots, N$, using the parameters $\theta^{(j)}$. Use the smoothed values to calculate S_{xb}, S_{bb}, S_{xx} given in (34)-(36).

4. Perform the M-Step. Update the estimates, $A, C, Q, R, \mu_0, \Sigma_0$ using (39)-(44).

5. Repeat Steps 2-4 to convergence.

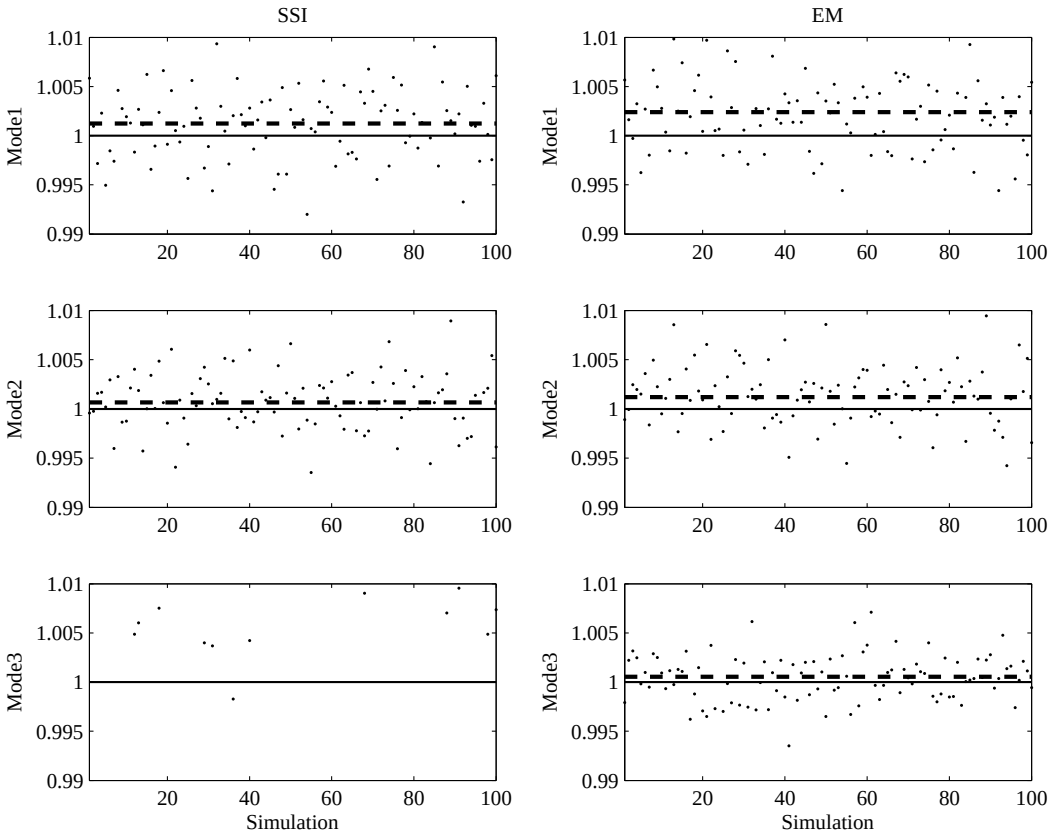


Figure 5: Eigenfrequency estimation results from 100 simulations. The estimates are divided by the true values (a value of 1 on the graphs indicates a perfect estimate). These relative frequencies are shown as dots. The scatter of this quantity gives an idea about the variance of the estimate. The average estimate is also shown (as a dashed line). The deviation of this quantity from 1 (full line) corresponds to the bias of the estimate. The rows show the modes; the columns represent the results of SSI and MLE+EM methods.

4 Numerical example

In this section, the effectiveness of the proposed identification method is evaluated via numerical simulations in the context of the ASCE benchmark problem for structural health monitoring.

4.1 Structural health monitoring benchmark problem

The benchmark studies currently consist of Phases I and II simulated and experimental problems. The benchmark structure is a four-story, two-bay by two-bay steel-frame scale model structure built in the Earthquake Engineering Research Laboratory at the University of British Columbia, Canada 2. The January 2004 issue of the Journal

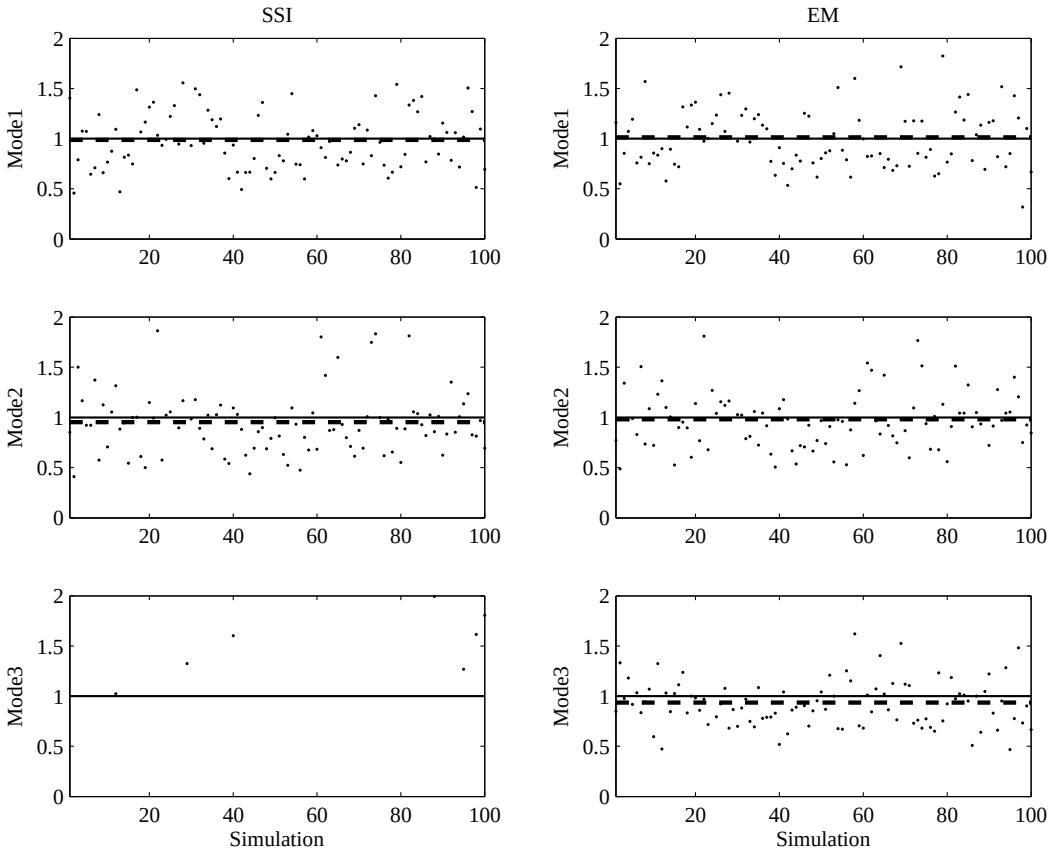


Figure 6: Damping ratio estimation results from 100 simulations. The estimates are divided by the true values (a value of 1 on the graphs indicates a perfect estimate). These relative damping ratios are shown as dots. The scatter of this quantity gives an idea about the variance of the estimate. The average estimate is also shown (as a dashed line). The deviation of this quantity from 1 (full line) corresponds to the bias of the estimate. The rows show the modes; the columns represent the results of SSI and MLE+EM methods.

of Engineering Mechanics contains the results of six different studies of the Phase I simulated benchmark problems, together with a definition and overview paper [7].

This paper focuses on Phase I of the analytical benchmark studies. A MATLAB-based finite element analysis code obtained from the IASC-ASCE SHM Task Group web site [8] is used to calculate the dynamic response of the prototype structure. Two finite-element models based on the actual test structure were developed to generate the simulated structural response data. The first is a 12DOF shear-building model that constrains all motion except two horizontal translations and one rotation per floor. The second is a 120DOF model that only requires floor nodes to have the same horizontal translation and in-plane rotation. The columns and floor beams are modeled as Euler-Bernoulli beams in both finite element models. A diagram of the analytical model

is shown in Fig. 2. The finite element models, by removing the stiffness of various elements, can simulate damage to the structure, and five damage patterns are defined for the structure. In this study has ben considered:

- 12 DOF undamaged structure.
- Classical damping (damping ratio equal to 0.01 for all modes).
- Sampling period is set as 0.001 s.
- Ten per cent root-mean-square (RMS) measurement noises.

The proposed method needs the following starting values (see Section 3.2.3):

- Starting values for the parameters $\theta^0 = (A, C, Q, R, \mu_0, \Sigma_0)$. In this work $\mu_0 = 0$, $\Sigma_0 = 0$; the state space parameters identified using Stochastic System Identification have been used as initial values for A, C, Q, R.
- 100 iterations for EM algorithm. Nevertheless, if

$$\frac{|L_{Y_N}(\theta^{j+1}) - L_{Y_N}(\theta^j)|}{|L_{Y_N}(\theta^j)|} < 10^{-10} \tag{45}$$

the iteration loop is stopped (L_{Y_N} is defined in Equation 19).

4.2 Discussion of results

Figure 3 shows the first three frequencies and mode shapes computed as the eigenvalues and eigenvectors of M and K matrices. So these are exact results. On the other hand, it has been simulated 100 cases, and each simulated result consists in accelerations at 16 points of the structure. Figure 2 shows the location of these points: four at each floor, 2 in x-direction and 2 in y-direction.

A 24 order state space model has been identified from each simulated data set using SSI method and the proposed MLE+EM method. The estimated eigenfrequencies for the first 20 simulations are plotted in Figure 4. In a general sense, the proposed method identifies accurate eigenfrecuencies for modes 3, 7 and 11. Mode 12 are not identified by both methods.

Finally, the modal parameters estimation results (eigenfrequencies, damping ratios and mode shapes) for the first three modes are represented in Figures 5 - 7. Modes 1 and 2 are well estimated with both methods, but mode 3 is only estimated with the proposed method.

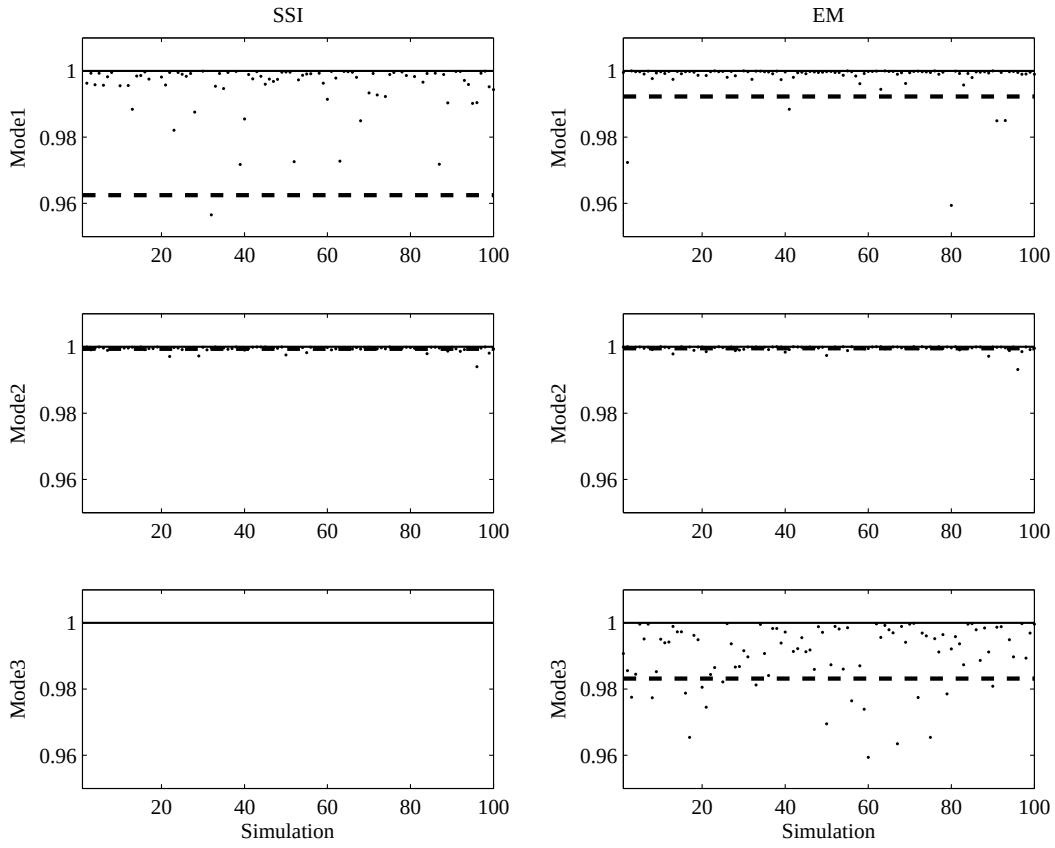


Figure 7: Mode shape estimation results from 100 simulations. The correlation between the estimated and the true modes shapes are shown (as dots). The average of the estimates is also shown (as a dashed line). The rows show the modes; the columns represent the results of SSI and MLE+EM methods.

5 Conclusions

This paper presents a time-domain stochastic system identification method based on Maximum Likelihood Estimation (MLE) and Expectation Maximization (EM) algorithm. Advantages of the proposed structural identification method can be summarized as follows: (i) the method is based on maximum likelihood, that implies minimum variance estimates; (ii) EM is a computational simpler estimation procedure than other optimization algorithms; (iii) estimate more parameters than SSI, and this estimates are accurate. On the contrary, the main disadvantages of the method are two: (i) EM algorithm is an iterative procedure and it consumes time until convergence is reached; (ii) this method needs starting values for the parameters.

The effectiveness of the proposed structural identification method has been evaluated through a numerical simulation study in the context of the ASCE benchmark problem. The numerical results show that the proposed method identifies eigenfre-

quencies, damping ratios and mode shapes reasonably well in the presence of 10% measurement noises even. These modal parameters are more acuratte than SSI estimated modal parameters.

References

[1] L. Ljung, "System Identification. Theory for the users". 2nd Ed., PTR Prentice-Hall, Upper Saddle River, N.J. 1999.

[2] P. Van Overschee, B. De Moor, "Subspace Identification for Linear Systems. Theory - Implementation - Applications". Kluwer Academic Publishers. 1996.

[3] R.H. Shumway, D.S. Stoffer, "Time series analysis and its applications". Springer. 2006.

[4] T. Katayama, "Subspace Methods for System Identification". Springer-Verlag London. 2005.

[5] M. Verhaegen, V. Verdult, "Filtering and System Identification. A least squares approach". Cambridge University Press. 2007.

[6] P. Van Overschee, B. De Moor. "Subspace algorithm for the stochastic identification problem", Automatica, 29, No. 3, 649-660. 1993.

[7] E. A. Johnson, H. F. Lam, L. S. Katafygiotis, J. L. Beck. "Phase I IASC - ASCE structural health monitoring benchmark problem using simulated data." J. Eng. Mech., 130(1), 315. 2004.

[8] IASC-ASCE SHM Task Group. Task group website (last visit on May 2010): http://bc029049.cityu.edu.hk/asce.shm/phase1/ascebenchmark.asp/

[9] B. Alicioglu, M. Lus, "Ambient vibration analysis with subspace methods and automated mode selection: case studies". Journal of Structural Engineering. ASCE. June 2008.

[10] B. Peeters, G. De Roeck, "Stochastic System Identification for Operational Modal analysis: A Review". Journal of Dynamic Systems, Measurement and Control. ASME. Vol. 123. December 2001.

Appendix

A Kalman filter

Property A.1 (The Kalman Filter) *For the state space model specified in* (3) *with initial conditions* $x_0^0 = \mu_0$ *and* $P_0^0 = \Sigma_0$*, for* $k = 1, 2, \ldots, N$*,*

$$x_k^{k-1} = A x_{k-1}^{k-1} \tag{46}$$

$$P_k^{k-1} = A P_{k-1}^{k-1} A^T + Q \tag{47}$$

with

$$x_k^k = x_k^{k-1} + K_k \epsilon_k \tag{48}$$

$$P_k^k = [I - K_k C] P_k^{k-1} \tag{49}$$

where

$$K_k = P_k^{k-1} C^T \Sigma_k^{-1} \tag{50}$$

$$\epsilon_k = y_k - \mathbf{E}\left[y_k | Y_{k-1}\right] = y_k - C x_k^{k1} \tag{51}$$

$$\Sigma_k = \mathbf{var}[C(x_k - x_k^{k-1} + v_k] = C P_k^{k-1} C^T + R \tag{52}$$

K_k is called the Kalman gain and ϵ_k are the innovations.

Property A.2 (The Kalman Smoother) *For the state space model specified in* (3) *with initial conditions x_N^N and P_N^N obtained via Property A.1, for $k = N, N-1, \ldots, 1$,*

$$x_{k-1}^N = x_{k-1}^{k-1} + J_{k-1}\left(x_k^N - x_k^{k-1}\right) \tag{53}$$

$$P_{k-1}^N = P_{k-1}^{k-1} + J_{k-1}\left(P_k^N - P_k^{k-1}\right) J_{k-1}^T \tag{54}$$

where

$$J_{k-1} = P_{k-1}^{k-1} A^T \left[P_k^{k-1}\right]^{-1} \tag{55}$$

Property A.3 (The Lag-One Covariance Smoother) *For the state space model specified in* (3)*, with K_k, J_k $(k = 1, 2, \ldots, N)$, and P_N^N obtained from Properties A.1 and A.2, with initial condition*

$$P_{N,N-1}^N = (I - K_N C) A P_{N-1}^{N-1} \tag{56}$$

for $k = N, N-1, \ldots, 2$

$$P_{k-1.k-2}^N = P_{k-1}^{k-1} J_{k-2}^T + J_{k-1}\left(P_{k,k-1}^N - A P_{k-1}^{k-1}\right) J_{k-2}^T \tag{57}$$

The demostration of the above properties can be found in [3].